

**REMARKS:**

This paper is herewith filed in response to the Examiner's final Office Action mailed on April 12, 2007 for the above-captioned U.S. Patent Application. This office action is a final rejection of claims 1-46 of the application.

More specifically, the Examiner has rejected:

claims 1, 4, 7, 9-15, 17, 19, 22, 25, 27-33, 35, 37, and 40 under 35 USC 102(b) as anticipated by Garfield;

claims 2-3, 6, 20-21, 24, 38-39, and 42 under 35 USC 103(a) as obvious over Garfield and Hull;

claims 18 and 36 under 35 USC 103(a) as obvious over Garfield and Kemp;

claims 5, 16, 23, 34, and 41 under 35 USC 103(a) as obvious over Garfield and Dittmar;

claims 8 and 26 under 35 USC 103(a) as obvious over Garfield and Drefahl;

claim 43 under 35 USC 103(a) as obvious over Garfield and Shivaratri;

claim 44 under 35 USC 103(a) as obvious over Garfield, Shivaratri, and Leiter; and

claims 45 and 46 under 35 USC 103(a) as obvious over Garfield, Shivaratri, and Drefahl.

The Applicants respectfully traverse the rejections.

Regarding the rejection of claim 1 the Examiner states;

“Garfield teaches a method [...] of **processing documents, comprising: partitioning document text and assigning semantic meaning to words** (p. 454, Objectives of linguistic analysis, sent. 4 and p. 490, 3rd para., sent. I), where assigning comprises applying a plurality of regular expressions (p. 469, 2nd-3rd para.),”

Claim 1 recites in part:

“A method to process a document, comprising: **partitioning document text and assigning semantic meaning to words of the partitioned document text**, where assigning comprises applying a plurality of regular expressions, rules and a plurality of dictionaries **to recognize chemical name fragments;**”

As cited Garfield discloses:

“Since modern formal linguistics certainly helps one to perceive semantic as well as grammatical categories more directly than the older, more intuitive methods, (comparable to a priori elucidation of chemical classifications) then it is of interest to explore the possibilities of using formal structural linguistics in studying the problem of chemical information retrieval,” (page 454, lines 16-24), and

“**Thus, the assignment of ‘meaning’ is conditioned by the syntactic methods that are employed for analyzing the chemical name and for generating the correct molecular formula** However, once the new approach is chosen, one must analyze each morpheme a little more closely. It is not sufficient, to know that nitro is NO<sub>2</sub>. It is necessary to learn that it is one nitrogen atom attached to two oxygen atoms, in which, one of the attachments is by a double bond.,” (page 490, lines 15-21).

In the prior office action the Applicants argued:

“Garfield discloses that the “the entire chemical name is punched on an IBM card or typed directly on a Unityper typewriter,” (page 485, second full para). Further, Garfield discloses that “working from right to left each character in the name is brought into the computer register one at a time and processed one at a time,” and “the character in process at any instant is referred to as the current character,” (page 485, second full para). Further, in Garfield, “each character (of the chemical name) is processed.” Garfield reads characters and groups of characters from the punch card until a dictionary match is found (page 486-488). The Examiner apparently considers Garfield’s punch card as anticipating the claimed document. **Because Garfield neither extracts identifying information from its punch cards, nor stores extracted punch card identifying information in association with anything else in a searchable index**, claim 1 is seen to be novel over Garfield.”

However, in the “Response to arguments” section of the Office Action the Examiner states:

**“Applicant points to the disclosure of Garfield in which chemical names are entered via punch cards** to be converted into chemical formulae. However, the use of punch cards is merely an illustrative proof of concept of Garfield’s teaching. On page 459, **Garfield introduces the concept of mechanical reading devices, ie optical character recognition, which would “avoid the costly step of manually creating a computer input”, i.e. punch cards (Mechanical reading device).** The mechanical reading device is then used to copy words (identifying information) from documents for indexing (**p. 460, lines 1-4**).” (emphasis added).

Further, in the rejection of claim 1 the Examiner states:

**“With respect to the amended additional limitations of extracting identifying information and storing extracted identifying information with chemical structural information, Garfield teaches a reading unit that selectively copies words for indexing and other purposes is seen to read on these limitations (**p. 460, lines 1-4**).”** (emphasis added).

Claim 1 recites in relevant part:

**“extracting identifying information from the document; and storing the extracted identifying information in association with the determined structural connectivity information in a searchable index.”**

The Applicant respectfully notes that based upon the Examiner’s response to the arguments and the statement in the rejection the Examiner appears to indicate that the “Mechanical reading device” or “a reading unit” disclosed in Garfield on page 460, lines 1-4 at least anticipates as claim 1 recites:

**“A method to process a document, comprising: partitioning document text and assigning semantic meaning to words of the partitioned document text, where assigning comprises applying a plurality of regular expressions, rules and a plurality of dictionaries to recognize chemical name fragments; recognizing any substructures present in the chemical name fragments; determining structural connectivity information of the chemical name fragments and recognized substructures; extracting identifying information from the document; and storing the extracted identifying information in association with the determined structural connectivity information in a searchable index.”**

Here Garfield actually discloses:

**“Nevertheless, a proto-type “reading” unit for selectively copying words for indexing and other purposes has been invented and built by this writer and is called the COPYWRITER (eg. Fourth Annual Report, Council on Library Resources, Washington, 1961, p. 30),” (emphasis added), (page 460, lines 1-4)**

However, prefacing this disclosure Garfield discloses:

**“In the work of indexing for the Index Chemicus, chemists must underscore pertinent chemical names and formulas. At present, there is no device available which would permit one to selectively “read” or “sense” printed texts, though the character recognition problem is gradually finding a solution,” (emphasis added), (page 459, fourth par. 1<sup>st</sup> and 2<sup>nd</sup> sentences).**

Clearly, this statement verifies that the “proto-type “reading” unit for selectively **copying words**” as disclosed by Garfield on page 460 and cited by the Examiner is not capable of reading or sensing printed texts. Further, in the invention there is no limitation that **chemical name fragments** or **identifying information** “must” be **underscored** in the document text as it appears is required by Garfield. Thus, the Applicants contend that for at least this reason the “proto-type “reading” unit for selectively **copying words**” in Garfield is also not capable of **“recognizing,” “partitioning” or “extracting”** from document text comprising chemical name fragments and identifying information as in claim 1.

Additionally, following this statement Garfield also discloses **“Assuming now that we have obtained some form of machine input either by character recognition or by manually creating a record in machine language, what do we wish to have done with this information?”** (page 460, lines 12-14). The Applicants contend that Garfield is merely assuming a capability which Garfield does not disclose but which is still used by the Examiner as relating to partitioning text, recognizing chemical name fragments, and extracting identifying information in a document as in the invention. Clearly, this assumption is not sufficient for the rejection of claim 1 under 35 USC 102(b).

The Applicants contend that the rejection which relies on a mere unsupported mention of a prototype application as anticipating claim 1 must fail. Garfield may recognize a chemical name entered on a punch card, which is not agreed with, but Garfield provides no details as to:

**“partitioning document text** and assigning semantic meaning to words of the partitioned document text, where assigning comprises applying a plurality of regular expressions, rules and a plurality of dictionaries **to recognize chemical name fragments,”** or

**“extracting identifying information** from the document; and **storing the extracted identifying information** in association with the determined structural connectivity information in a searchable index,” as in claim 1.

Imputing the **concept of partitioning, recognizing, and extracting text of a document** approach to Garfield is clearly hindsight. A 35 USC 102 rejection requires that the cited art **disclose to the specificity of the rejected claim**; *Verve, LLC v. Crane Cams, Inc.*, 311 F.3d 1116, 1120, 65 USPQ2d 1051 (Fed. Cir. 2002) (“**A single reference must describe the claimed invention with sufficient precision and detail to establish that the subject matter existed in the prior art**”). It is axiomatic that a 35 USC 102(b) rejection requires strict identity with every claim element. For at least the reasons stated above Garfield fails to anticipate claim 1, and the rejection of claim 1 should be removed.

Further, as the independent claims 19, and 37 recite language similar to claim 1 as stated above, the rejection should be removed for all the claims 1, 19, and 37.

In regards to the rejection of claims 4, 22, and 40 the Examiner states:

“Regarding claims 4, 22 and 40, Garfield teaches the searching the index by fragment or substructure name and/or connectivity (p. 466, Relationship Between Nomenclature and Searching).”

Further, in the Response to arguments section the Examiner states:

“Garfield stores the chemical information (morpheme) in a “morpheme storage

area along with its appropriate meaning” reading on storing structural connectivity information in association with identifying information (p. 486, dictionary math routine” lines 2-4).”

Here Garfield discloses:

“When the computer analysis of the chemical name is completed, the parsed expression that results from the analysis could be used by the computer to perform very adequate generic as well as specific searches. If the chemists specifies the type of chemical in which he is interested in terms of morphemes instead of conventional chemical class names, generic searches become quite simple,” (page 466, lines 28-32); and

“The dictionary match routine will compare the contents of the alpha storage with the dictionary and will find a match for ene. Since this morpheme is not on the pent-oct list the morpheme ene will be placed in a special calculation and morphene storage area along with it’s appropriate meaning.” (page 486, lines 21-24).

Claim 4 recites:

A method as in claim 1, further comprising **searching the index by a combination of at least one of fragment and substructure name, and at least one of fragment and substructure connectivity.**

The Applicants contend that Garfield merely mentions generic as well as specific searches. However, Garfield does not disclose wherein “generic as well as specific searches” is performed “**by a combination of at least one of fragment and substructure name, and at least one of fragment and substructure connectivity,**” as in claim 4. Thus, the rejection of claim 4 should be removed. Further, as claims 22 and 40 recite language similar to claim 4 as stated above the rejection should be removed for all the claims 4, 22, and 40.

Regarding the rejection of claims 10-11 and 28-29 the Examiner states:

“Garfield teaches a dictionary of stop words (p. 487, 2nd para, 2nd to last sentence).”

Here Garfield discloses:

“Determining whether a character is ignorable is done by a dictionary sub-routine, in which the computer compares each current character with a complete list of ignorable characters **consisting of the integers 1 to 8, hyphen, comma, prime, and colon,**” (page 487, 2nd para, 2nd to last sentence).

Here it is clear to the Applicants that Garfield does not disclose “where said plurality of

dictionaries comprise a dictionary of **stop words** to eliminate erroneous **chemical name fragments**,” as in claim 10 or “**filtering recognized chemical name fragments** using a list of **stop words** to eliminate erroneous chemical name fragments,” as in claim 11. Garfield merely lists “ignorable characters **consisting of the integers 1 to 8, hyphen, comma, prime, and colon.**” Thus, the rejection of claims 10 and 11 should be removed. Further, as the claims 28 and 29 recite similar language of claims 10 and 11 respectively, the rejection should be removed for all the claims 10, 11, 28, and 29.

Further, regarding the rejection of claims 13-15 and 31-33 the Examiner states:

“Regarding claims 13-15 and 31-33, Garfield teaches the application of **regular expressions comprised of a plurality of patterns** (further comprised of characters numbers and punctuation; cl.14 (**p. 487, 2nd para., sent. 5**)) in which punctuation characters are maintained or removed (cl.13) (Dictionary match routine, p.486) and where punctuation can be at least one of parenthesis, square bracket, hyphen, colon and semicolon (cl. 15) (**p. 487, 2nd para., sent. 5, “paren” and next sentence, “hyphen”**).”

The Applicants note that “ p. 487, 2nd para., sent. 5[...] and next sentence” as cited by the Examiner in the rejection of claims 13-15 and 31-33 appears to reference “p. 487, 2nd para, 2nd to last sentence” as cited by the Examiner above in the rejection of claim 4.

Thus, here and as stated above Garfield discloses:

“Determining whether a character is ignorable is done by a dictionary sub-routine, in which the computer compares **each current character** with a complete list of ignorable characters **consisting of the integers 1 to 8, hyphen, comma, prime, and colon,**”

The Applicants contend that in Garfield comparing “each current character” with a list of ignorable characters consisting of “integers 1 to 8, hyphen, comma, prime, and colon” does not disclose “where application of said **regular expressions and rules** results in punctuation characters being one of maintained or removed between chemical name fragments **as a function of context,**” as claim 13 recites in part.

Further, for at least the reasons stated Garfield is not seen to disclose as claim 14 recites in part:

“where said **regular expressions comprise a plurality of patterns**, individual ones of which are comprised of at least one of characters, numbers and punctuation.”

Moreover, for at least the reasons stated Garfield is not seen to disclose as claim 15 recites in part:

“where the punctuation comprises at **least one of parenthesis, square bracket, hyphen, colon and semi-colon.**”

Further, for at least the reasons that the claims 4, 7, 9-15, and 17; claims 22, 25, 27-33, and 35; and claim 40 depend from claims 1, 19, and 37 respectively the rejections should be removed for all the claims 1, 4, 7, 9-15, 17, 19, 22, 25, 27-33, 35, 37, and 40.

Regarding the rejection of claim 6 under 35 USC 103(a) the Applicants respectfully disagree with the rejection.

The Examiner admits:

“Garfield does not teach storing keywords and identifying information in association with structural connectivity and then searching an index with a keyword and a fragment name or connectivity.”

Further, the Examiner states:

“**Hull teaches extracting keywords from the document (col. 9, lines 15 -32).** Extracted identifying information is stored in association with structural connectivity information in a searchable matrix (index) (col.10, lines 32-52).”

Hull actually discloses:

“Referring to FIG. 3, in step S300, **a user and/or a computer generates or creates chemical and textual descriptors** for each compound represented in the database. The textual descriptors may, for example, originate from a collection of documents, or other text source,



in, say, ASCII format or other suitable format,” (col. 9, lines 18-23).

Thus, the Applicants contend that the Examiner’s statement that “Hull teaches extracting keywords from the document” is incorrect. Hull is not seen as “extracting keywords” but rather as Hull discloses **a user and/or computer generates or creates chemical descriptors**. Thus, Hull is not seen to disclose or suggest as claim 6 recites in part:

“A method as in claim 1, wherein **extracting** further comprises **extracting keywords from the document** and wherein storing comprises **storing the extracted identifying information and the extracted keywords** in association with the determined structural connectivity information in the searchable index”

Further, as claims 24 and 42 recite language similar to claim 6 as stated above, Hull is not seen to disclose or suggest all the claims 6, 24, and 42.

Moreover, the Applicants contend that even if Garfield were modified in view of Hull, which the Applicants do not agree is feasible or possible, for at least the reasons already stated the result would still not disclose or suggest the claims. Therefore, the rejection of the claims should be removed.

Regarding the rejection of claims 18 and 36 over Garfield in view of Kemp the Applicants disagree with the rejection.

Although the Applicants do not agree with the Examiner’s characterization that “Kemp et al. teach the tokenization of documents into a sequence of tokens (p. 547, 2nd para, sent. 2)” as stated in the rejection, the Applicants contend that Kemp still is not seen to address the shortfalls of Garfield as stated above. Thus, the Applicants contend that even if Garfield were modified in view of Kemp, which the Applicants do not agree is feasible or possible, for at least the reasons already stated the result would still not disclose or suggest the claims. Therefore, the rejection of the claims should be removed.

Furthermore, regarding the rejection of claims 5, 16, 23, 34, and 41 the Examiner admits:

“Garfield does not teach searching an index by at least one of a fragment or

substructure connectivity using a graphical user interface or characters comprising at least one of upper case C, O, R, N, H.”

Further, the Examiner states:

“Dittmar et al. teach searching an index by at least one of a fragment or substructure connectivity (p.99, col. 2, para2, sent.1) using a graphical user interface (p. 93, col. 1, para. 3, sent. 2).

Dittmar et al. teach or characters comprising at least one of upper case C, O, R, N, H (p. 98, cal. 1 par 2, sent. 3; para.3, sent. 1; and p. 99, para 2-3).

It would have been obvious to combine the teaching Garfield with the teach Dittmar et al. because Dittmar et al. teach implementation of a user interface to simplify searching (p. 93, cal. 1, para 3, sent. 1).”

The Applicants note that based upon the Examiner’s identifiers the Applicants can not find wherein the reference discloses the claims as stated by the Examiner. The Applicants respectfully request clarification in a non-final Office Action or removal of the rejection of the claims.

In addition, although the Applicants do not agree with the Examiner’s characterization of Dittmar as stated in the rejection, the Applicants contend that Dittmar still is not seen to address the shortfalls of Garfield as stated above. Thus, the Applicants contend that even if Garfield were modified in view of Dittmar, which the Applicants do not agree is feasible or possible, for at least the reasons already stated the result would still not disclose or suggest the claims. Therefore, the rejection of the claims should be removed.

Regarding the rejection of claims 8 and 26 over Garfield in view of Drefahl; the rejection of claim 43 over Garfield in view of Shivartri; the rejection of claim 44 over Garfield in view of Shivartri and in further view of Leiter; and the rejection of claims 45 and 46 over Garfield in view of Shivartri and in further view of Drefahl the Applicants respectfully traverse the rejections.

For at least the reasons already stated the Applicants contend that neither Drefahl nor Shivartri

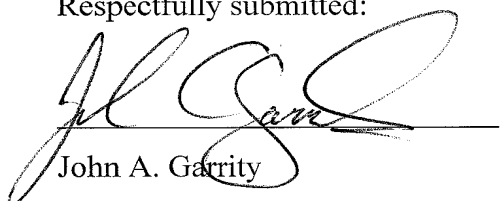
S.N.: 10/797,359  
Art Unit: 1631

nor Leiter is seen to overcome the shortfall of Garfield as stated above. Further, even if Garfield were modified in view of Drefahl or Shivartri or Leiter, which is not agreed is feasible or possible, for at least the reasons stated the combination would still not disclose or suggest the claims. Thus, the rejection of these claims should be removed.

Based on the above explanations and arguments, it is clear that the references cited cannot be seen to anticipate claims 1-46. The Examiner is respectfully requested to reconsider and remove the rejections of claims 1-46 and to allow all of the pending claims 1-46.

For all of the foregoing reasons, it is respectfully submitted that all of the claims now present in the application are clearly novel and patentable over the prior art of record. Should any unresolved issue remain, the Examiner is invited to call Applicants' attorney at the telephone number indicated below.

Respectfully submitted:

  
John A. Garrity

6/11/07  
Date

Reg. No.: 60,470

Customer No.: 48237

HARRINGTON & SMITH, PC

4 Research Drive

Shelton, CT 06484-6212

Telephone: (203)925-9400

Facsimile: (203)944-0245

email: jgarrity@hspatent.com

### **CERTIFICATE OF MAILING**

I hereby certify that this correspondence is being deposited with the United States Postal Service

S.N.: 10/797,359  
Art Unit: 1631

as first class mail in an envelope addressed to: Commissioner for Patents, P.O. BOX 1450,  
Alexandria, VA 22313-1450.

---

Date

---

Name of Person Making Deposit